

## Research

## Matrix attachment regions as targets for retroviral integration

Chassidy N Johnson and Laura S Levy\*

Address: Department of Microbiology & Immunology and Tulane Cancer Center, Tulane University School of Medicine, New Orleans, Louisiana, 70112, USA

Email: Chassidy N Johnson - [cjohnso9@tulane.edu](mailto:cjohnso9@tulane.edu); Laura S Levy\* - [llevy@tulane.edu](mailto:llevy@tulane.edu)

\* Corresponding author

Published: 19 August 2005

Received: 13 June 2005

*Virology Journal* 2005, **2**:68 doi:10.1186/1743-422X-2-68

Accepted: 19 August 2005

This article is available from: <http://www.virologyj.com/content/2/1/68>

© 2005 Johnson and Levy; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** The randomness of retroviral integration has been debated for many years. Recent evidence indicates that integration site selection is not random, and that it is influenced by both viral and cellular factors. To study the role of DNA structure in site selection, retroviral integration near matrix attachment regions (MARs) was analyzed for three different groups of retroviruses. The objective was to assess whether integration near MARs may be a factor for integration site selection.

**Results:** Results indicated that MLV, SL3-3 MuLV, HIV-1 and HTLV-1 integrate preferentially near MARs, specifically within 2-kilobases (kb). In addition, a preferential position and orientation relative to the adjacent MAR was observed for each virus. Further analysis of SL3-3 MuLV insertions in common integration sites (CISs) demonstrated a higher frequency of integration near MARs and an orientation preference that was not observed for integrations outside CISs.

**Conclusion:** These findings contribute to a growing body of evidence indicating that retroviral integration is not random, that MARs influence integration site selection for some retroviruses, and that integration near MARs may have a role in the insertional activation of oncogenes by gammaretroviruses.

### Background

An essential step in the replication cycle of all retroviruses is integration of the double-stranded DNA proviral form of the genome into host DNA. The degree of randomness of proviral integration has been debated for many years [1,2]. Studies have suggested that DNaseI hypersensitive sites [3-7], AT-rich regions [8], transcriptionally active regions [2,9-12], repeat elements including Alu and LINE elements [13] and regions of DNA bending, specifically regions with the most DNA distortion [14-18], are preferred sites of proviral integration. Alternatively, studies have shown that high levels of transcription disfavor integration of avian leukosis virus (ALV) [2]. The conflicting

results that have been reported may be explained by the small sample sizes examined or by potential biases introduced from the cloning strategies used to identify insertion sites. In addition, many of the studies were performed *in vitro*, and thus did not take into account the native conformation of chromatin. Before the completion and publication of the human and mouse genome databases, theories for randomness of retroviral integration were difficult to prove or disprove because of the technical challenge of analyzing a large sample size of integrations from infected cells. Since publication of the genome databases, several studies have isolated and mapped hundreds of proviral insertion sites for murine leukemia virus

(MLV), human immunodeficiency virus type-1 (HIV-1), avian sarcoma virus (ASV) and human T-cell leukemia virus type-1 (HTLV-1) [11,12,19,20]. For those viruses, the results showed preferential integration into transcriptionally active NCBI Reference Sequences (RefSeqs), but distinct patterns of integration were evident as well. These studies provided strong evidence that distinct viruses differ in proviral integration patterns, but that integration is clearly non-random. The specific pressures that influence site selection for retroviral integration remain incompletely understood.

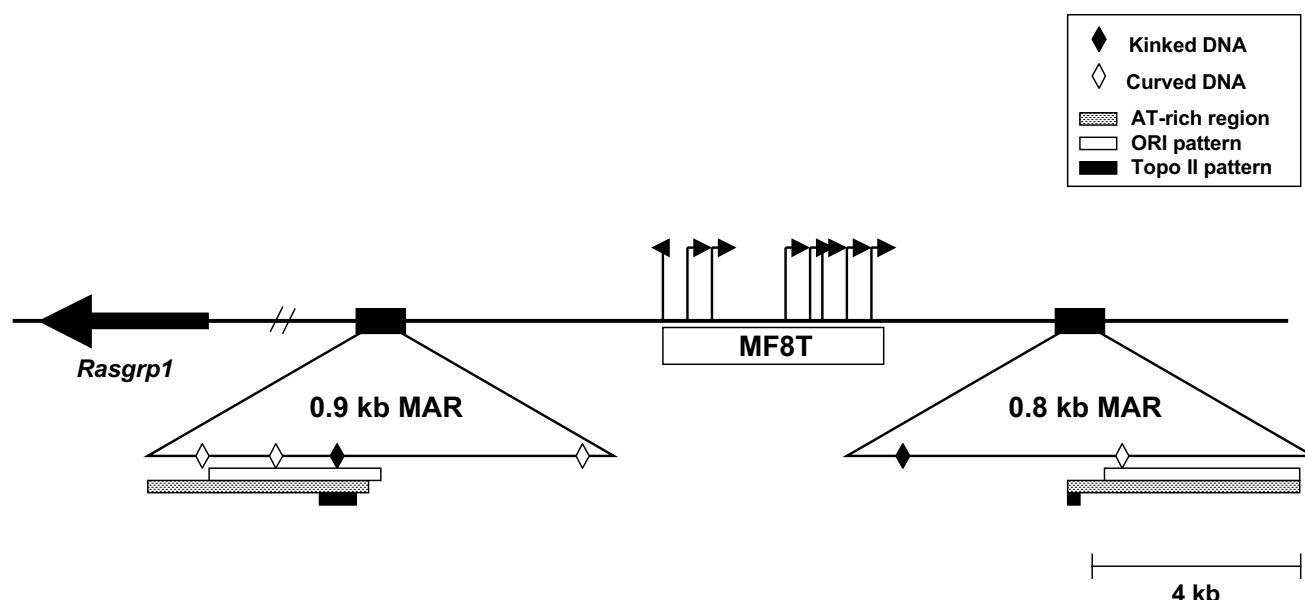
Accumulating evidence indicates that retroviral integration site selection is influenced by properties of cellular DNA structure [11,21-24]. A recent large-scale study found that DNA structural features such as bendability and A-philicity served as preferred integration sites [22]. The present study was performed to assess the role of matrix attachment regions (MARs) in retroviral integration site selection. MARs are DNA sequences located at the bases of DNA loops that attach to the nuclear matrix, and are thus positioned near the machinery for DNA replication, transcription, RNA processing and transport (reviewed in [25]). There is no consensus sequence that defines a MAR; however, MARs are commonly found to have intrinsic DNA bending properties, to contain transcription factor binding sites, AT-rich stretches, sites for topoisomerase I and II binding and cleavage, and high unwinding potential [26,27]. MARs function as structural regulatory elements by organizing the DNA into loop domains. Studies have shown that MARs influence the expression of cellular genes, and can enhance viral gene expression when in the vicinity of viral promoters and enhancers [28-30]. This property has made the inclusion of MARs in gene therapy vectors attractive for enhanced and prolonged expression of the transgene in a specific cell-type or developmental stage [31-33]. MARs have been implicated in virus-mediated malignancies, particularly as targets of integration by small DNA tumor viruses. Specifically, integrated SV40, HBV, HPV16 and HPV18 have been found within or in close proximity to MARs in tumors or transformed cell lines [34]. Other reports indicate that HTLV-1 and HIV-1 may integrate preferentially near MARs [34,35].

The gammaretroviruses represent a group of mammalian oncogenic retroviruses typically associated with the induction of long-latency leukemia and lymphoma in the natural host. Gammaretroviruses do not encode an oncogene or any other gene to which their malignant potential can be directly attributed. Rather, their ability to induce tumors has been linked to a process termed insertional activation, in which integration of the proviral genome into host DNA is associated with activated expression of an adjacent oncogene. When the same genetic locus is

observed to be interrupted by proviral integration in multiple independent tumors, it is inferred that the commonly interrupted locus encodes an oncogene whose activation is relevant to tumor induction [36-38]. Such a locus is referred to as a common insertion site (CIS). We recently described CISs utilized by a recombinant gamma-retrovirus, MoFe2-MuLV (MoFe2), in T-cell lymphomas in the NIH/Swiss mouse. To construct MoFe2, the U3 region of the Moloney murine leukemia virus (M-MuLV) long terminal repeat (LTR) was substituted with homologous sequences from a natural isolate of feline leukemia virus termed FeLV-945 [39]. FeLV-945 is characterized by a unique motif in the U3 region of the LTR, which contains a single copy of the transcriptional enhancer followed downstream by the tandem triplication of a 21-bp sequence. Substitution of FeLV-945 LTR sequences into M-MuLV was shown to alter the pattern of insertional activation and to identify new CISs [40]. As described below, the identification of two potential MARs near a CIS in MoFe2-induced lymphomas suggested that MARs may represent a determinant of integration site selection. That hypothesis was addressed in the present study by analyzing the proximity of proviral integrations to MARs in lymphomas and in unselected cultured cells. The patterns of integration with respect to MARs were compared for three groups of retroviruses, including several murine gamma-retroviruses, human deltaretrovirus (HTLV-1) and lentivirus (HIV-1).

## Results

Previous studies showed that inoculation of neonatal mice with MoFe2 resulted in the development of T-cell lymphoma. Analysis of patterns of common proviral insertion in lymphomas revealed that MoFe2 utilized a set of CISs distinct from either parent virus from which it was constructed [39,40]. Sequence surrounding one of the previously described CISs in MoFe2-induced lymphomas, termed *MF8T* (*Rasgrp1*), was analyzed for the presence of MARs using a MAR-prediction program termed MAR-Finder <http://futuresoft.org>. MAR-Finder is a statistical algorithm that analyzes the pattern density for characteristic DNA sequence motifs that predict the occurrence of MARs, including replication origins, TG-richness, curved DNA, kinked DNA, topoisomerase II recognition and cleavage sites and AT-richness. MAR-Finder has been previously validated for predicting the presence of MARs [34,41,42]. An alternative method to predict MARs is based on detecting the location and extent of stress-induced duplex destabilization (SIDDD) through the use of a statistical algorithm termed WebSIDDD [43-45]. Although this method has been validated to predict the presence of MARs accurately, recent evidence indicates that stress-induced destabilization of duplex DNA is not sufficient for a sequence to bind to the nuclear matrix; thus, the use of SIDDD for the prediction of MARs may lead

**Figure 1**

Physical map of the MF8T locus. Depicted is the 3.9-kb domain of common proviral insertion designated MF8T. Vertical lines represent the positions of the proviral integrations with the transcriptional orientation of provirus depicted by the direction of the arrow. Depicted is *Rasgrp1*, the predicted oncogene in the MF8T locus. Two predicted MARs of 0.9-kb and 0.8-kb in size are located 5.1-kb and 3.6-kb from the domain of common insertion. Also depicted are structural motifs typical of MARs, including kinked DNA, curved DNA, AT-rich regions, ORI patterns and Topoisomerase II cleavage site patterns.

to false positives [46]. Using MAR-Finder, the results indicated the presence of two MARs in the 60-kb sequence surrounding *MF8T*, located 5.1-kb and 3.6-kb from the domain of common insertion (Figure 1). The predicted elements were observed to be enriched in motifs characteristic of MARS, including kinked DNA, curved DNA, AT-rich regions, origin of replication patterns and vertebrate and *Drosophila* topoisomerase II consensus sequences [26,27]. The close proximity of two MARS to the *MF8T* CIS suggested that integration near MARS may represent a mechanism for retroviral target site selection. To evaluate this possibility, the distance from proviral integration to predicted MARs was analyzed for three different groups of retroviruses, specifically murine gammaretroviruses (MoFe2, SL3-3 MuLV, MLV) human deltaretrovirus (HTLV-1) and lentivirus (HIV-1). Sequence information on MoFe2 integrations was obtained from the CISs and other insertion sites identified previously from a large collection of MoFe2-induced tumors [40]. MoFe2 integration sites were also analyzed from acutely infected SC-1 cells. In total, 42 MoFe2 integration sites were identified and analyzed in the present study. SL3-3 MuLV (SL3-3) inte-

gration sites had been previously identified from T-cell lymphomas in NIH-Swiss mice by inverse PCR [47]. In total, 86 SL3-3 integration sites were examined in the present study [47]. MLV and HIV-1 integration sites had been previously identified from HeLa cells infected with pseudotyped retroviral genomes [19]. From the 903 MLV and 379 HIV-1 insertions identified in that study, 49 (MLV) or 41 (HIV-1) integration sites for each virus were chosen at random for the present analysis. HTLV-1 integration sites from tumor-derived cells lines or from ATLL patients had been previously identified [8,12,34], 26 of which were examined in the present study. For each integration site examined in the present study, host-virus junction fragment sequences were obtained from GenBank <http://www.ncbi.nlm.nih.gov/Genbank/index.html> or the Mouse Retroviral Tagged Cancer Gene Database (RTCGD; <http://rtgd.ncifcrd.gov>) and the integration sites were thereby positioned in the respective mouse or human genome using the NCBI mouse or human genome database <http://www.ncbi.nlm.nih.gov/genome/guide/human/> or <http://www.ncbi.nlm.nih.gov/genome/guide/mouse/>.

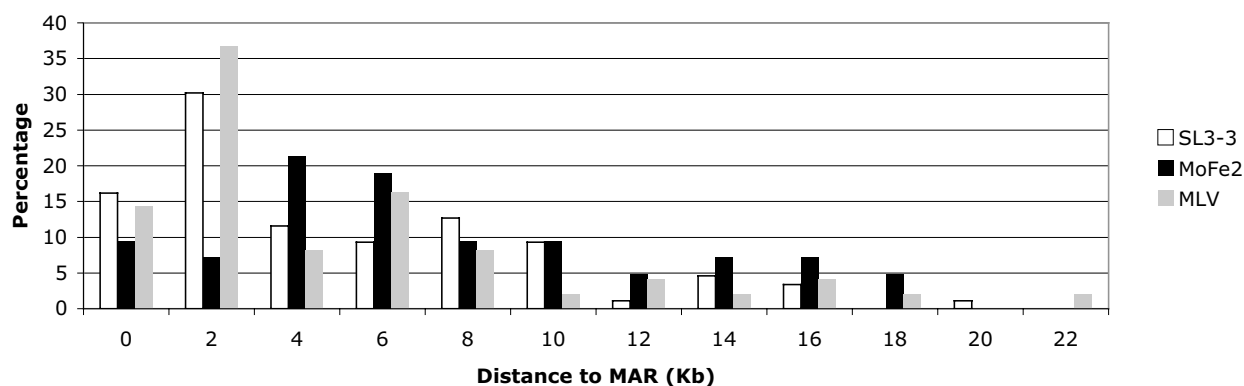
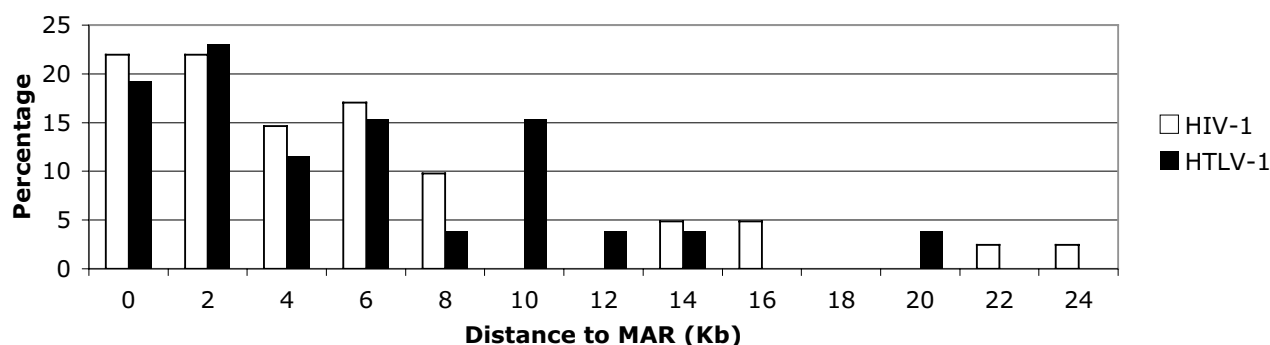
Initial analysis of insertion sites and their proximity to MARs revealed that some integrations were located more than 20-kb from a predicted MAR; therefore, to ensure a thorough identification of MARS in the vicinity of proviral integrations, 60-kb of sequence information surrounding each insertion site was obtained from the respective genome for analysis. Using 60-kb of sequence information surrounding each integration event, the distance from the proviral insertion site to the closest predicted MAR was plotted as the percentage of integration events analyzed (Figure 2). For the murine gammaretroviruses, the results indicated a preference to integrate within 2-kb of a predicted MAR. For example, 46% of SL3-3 integrations and 50% of MLV integrations occurred within 2-kb of a predicted MAR (Figure 2A). It has been reported that MARs occur every 10-kb in the mammalian genome [34,41]. Based on this report, a Monte Carlo simulation was performed where the mean distance to the closest MAR was computed under the assumption that viral integration occurs randomly with respect to regions that are predicted MARs and that MARs occur every 10-kb. The results indicated that, under these assumptions, the mean distance to the closest MAR during a random integration event would be 4-kb [34]. Thus, preferential integration near MARs is indicated for SL3-3 and MLV. By comparison, MoFe2 integration did not show the same preference (Figure 2A); rather, the distribution of MoFe2 integration sites in relation to MARs was significantly different from the distribution observed for SL3-3 and MLV ( $p < 0.01$ ). In fact, the distribution of MoFe2 insertions in relation to MARs was consistent with the expectation for random integration. The same analysis was then performed on HTLV-1 and HIV-1 to determine if integration near MARs is also common for retroviruses that do not act in disease induction by insertional activation. The results indicated a preference for integration near MARs, since 43.9% of HIV-1 integrations and 42.3% of HTLV-1 integrations occurred within 2-kb of a predicted MAR (Figure 2B). As expected, a small percentage of integration events occurred more than 10-kb from a predicted MAR (Figure 2). In fact, for some integrations sites, the closest MAR in one direction was more than 60-kb away (data not shown). These results illustrate that, although MARs are predicted to be positioned at 10-kb intervals, there are regions of DNA that are either enriched or deficient in MARs as well.

Previous reports have indicated that MAR-mediated enhancement of viral gene expression is directional [32,34]. Other reports, in contrast, have indicated that MARs function to enhance gene expression in an orientation- and position- independent manner when located near the promoter [48]. To examine whether the preferred gammaretroviral integration near MARs is directional, it was next determined whether the closest predicted MAR

was located upstream or downstream of the proviral integration site with respect to the transcriptional direction of the genetic locus. Results of the analysis, plotted as a percentage of integration events, indicated that the majority of MLV integrations occurred 1- to 2-kb from a predicted MAR on the downstream side (Figure 3A). For SL3-3, it was useful to consider independently the integrations previously identified as CISs in tumor DNA, since those integrations presumably function to activate nearby oncogenes [47]. Interestingly, SL3-3 insertions identified as CISs were found to integrate commonly within 2-kb from a predicted MAR and to be positioned on the upstream side. Of 31 such insertions examined, 29% were integrated within 2-kb upstream as compared to 5.8% integrated within 2-kb downstream of a predicted MAR (Figure 3A). By comparison, 44 SL3-3 integrations identified as only single insertion sites (ISs) did not show the same directional preference for integration near MARs (Figure 3A). These findings imply that SL3-3 integration immediately upstream of MARs within CISs may be related to insertional activation of the adjacent oncogene.

When examined by the same approach, analysis of HIV-1 and HTLV-1 integrations indicated that the majority of proviral insertions occurred near MARs, and 80% of the HIV-1 proviral integrations that occurred within 1- to 2-kb of a MAR were positioned downstream (Figure 3B). HTLV-1, while integrated preferentially within 2-kb of a MAR, did not show a position preference. A recent study also analyzed HIV-1 integration sites for their proximity to MARs. Consistent with our findings, that study indicated HIV-1 integration near MARs, specifically in the downstream position [35]. Another study, however, reported that MARs are commonly found downstream from the sites of HTLV-1 integrations [34]. As noted, we did not observe a position preference for HTLV-1 integrations relative to MARs (Figure 3B). The conflicting results may be due to the small sample size ( $n = 3$ ) examined in the previous study.

Several recent studies have reported that HIV-1, MLV and HTLV-1 integrate preferentially into genes [12,19,20]. With these findings in mind, SL3-3 and MoFe2 insertion sites were analyzed to determine whether a preference is evident for integration into RefSeqs. The analysis revealed that 17.6% SL3-3 integrations at CISs, 40.3% of SL3-3 integrations at single insertion sites, and 33.3% MoFe2 insertions occurred within RefSeqs (data not shown). By comparison, the frequency of integration into genes by random chance has been estimated at 22% [12,19,20]. Thus, preferential integration into genes was identified for MoFe2 and SL3-3 at single insertion sites, although not for SL3-3 integrated at CISs. Analysis was then performed to determine if preferred integration into genes was associated with integration near MARs. Using the NCBI mouse

**A.****B.****Figure 2**

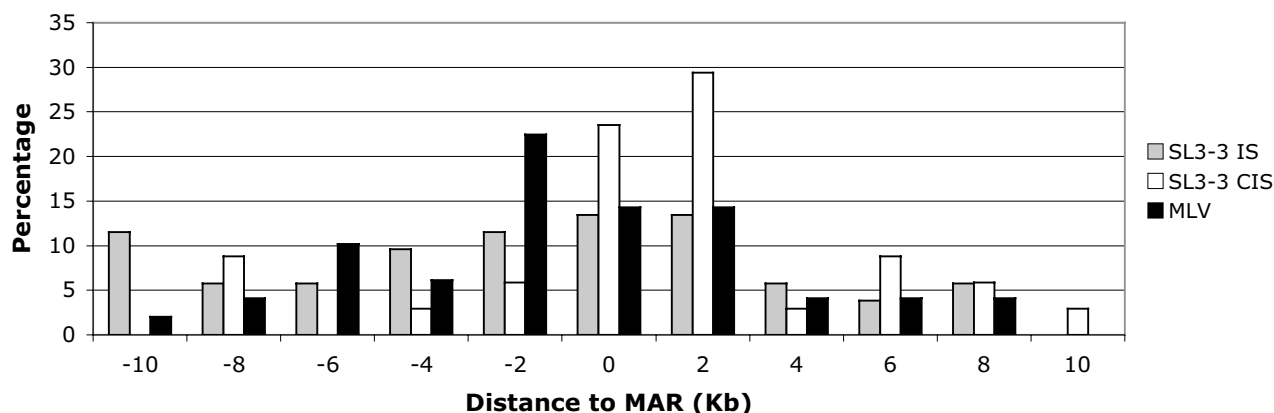
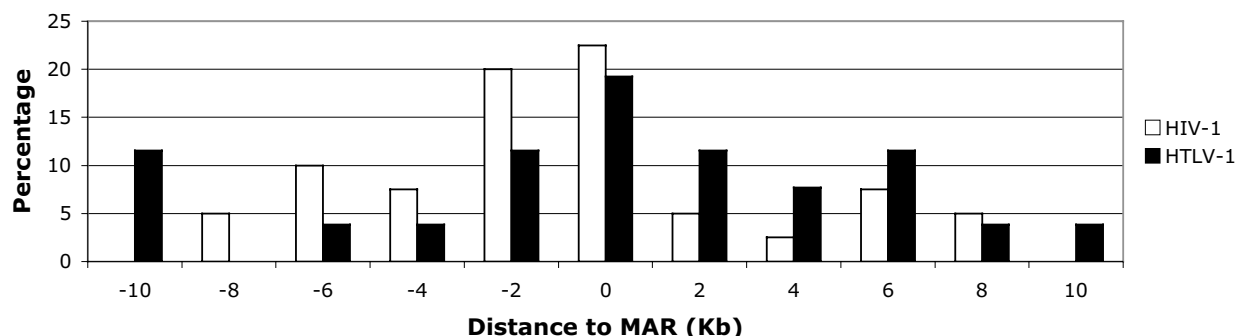
Distance of closest predicted MAR to proviral insertion site. Results are plotted as the percentage of integration events that occurred within 25-kb from a MAR using MAR-Finder for (A) gammaretroviruses (SL3-3, MoFe2 and MLV) and (B) HIV-1 and HTLV-1. SL3-3 and MLV integration distribution was significantly different than MoFe2 as determined by a one-way ANOVA followed by Tukey's multiple comparison test.

or human genome database, integration events were first grouped as to whether they occurred within or between genes. For each of the groups, the percentage of integrations that occurred within 2-kb of a predicted MAR was then determined (Figure 4). The results indicated no relationship to the nearest MAR when integration occurred within genes for SL3-3 at single insertion sites, MLV, MoFe2 or HTLV-1. In contrast, 71.4% of HIV-1 integrations that occurred within genes were observed to occur within 2-kb of a MAR. A strong relationship to MARs was also observed for SL3-3 integrations at CISs that occurred between genes. Of these integrations, 68.8% were observed to occur within 2-kb of a predicted MAR.

## Conclusion

Evidence is accumulating to indicate that proviral integration is not random, and that the secondary structure of

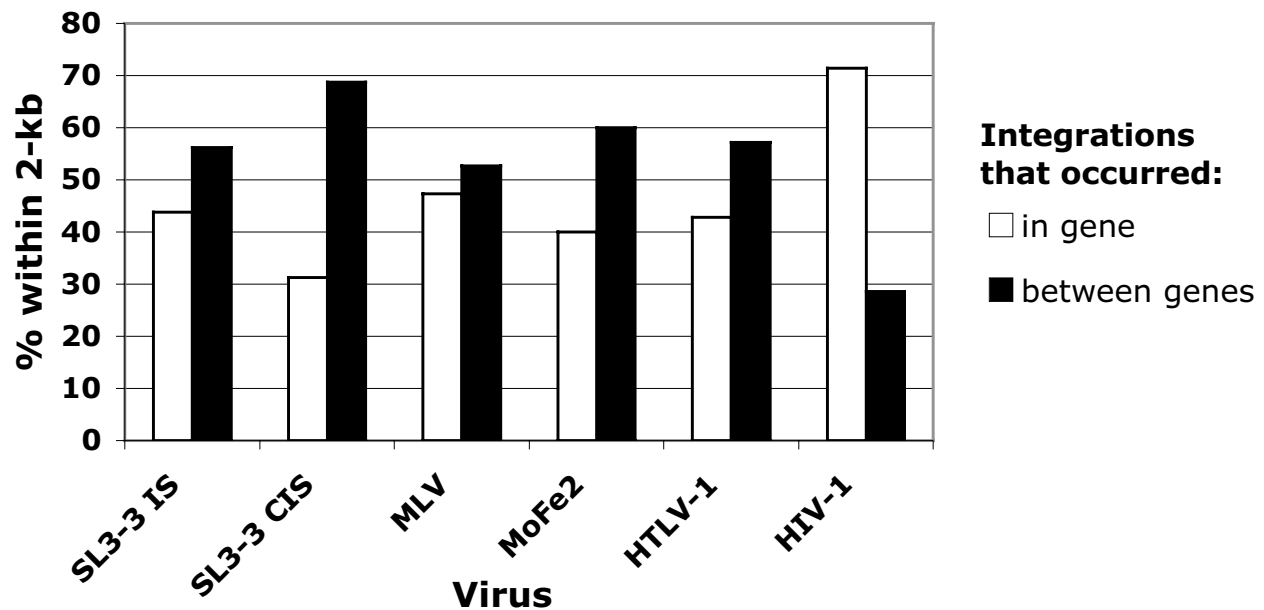
DNA plays a major role in integration site selection [2-18]. In the present study, the integration patterns of three different groups of retroviruses with distinct mechanisms of disease induction were analyzed to determine if integration near MARs is a common mechanism of retroviral integration site selection. The results indicated that gammaretroviruses (MLV and SL3-3), lentivirus (HIV-1) and deltaretrovirus (HTLV-1) integrate preferentially near MARs, specifically within 2-kb (Figure 2). These results suggest that integration near MARs is a common mechanism of retroviral integration site selection. The findings are consistent with the previous identification of preferred integration sites that contained sequence motifs such as DNaseI hypersensitive sites [3-7], AT-rich regions [8], transcriptionally active regions [2,9-12], and regions of DNA bending, specifically regions with the most DNA distortion [14-18], all of which are motifs shared by MARs. A

**A.****B.****Figure 3**

Position of MAR closest to the proviral integration site. The closest predicted MAR to the site of proviral insertion was determined to be located upstream or downstream from the site of insertion with respect to the transcriptional direction of the genetic locus. The results are plotted as the percentage of integrations that occurred up to 10-kb from a MAR for (A) SL3-3 insertions at single insertion site (SL3-3 IS), SL3-3 insertions at common insertion sites (SL3-3 CIS), MLV and (B) HTLV-1 and HIV-1.

recent study analyzed the proximity of retroviral integration to MARs when the virus was delivered to the cell by infection or by electroporation of naked DNA [49]. The results showed a strong correlation for integration near MARs during infection, but not when transfected as naked DNA. These results further support a role for MARs in integration site selection during retroviral infection. There are several possible explanations for preferential integration near MARs. One possibility is that MARs, due to their position at the bases of chromatin loops, are likely to be the first region of the DNA encountered by the provirus when entering the nucleus. A second possibility is that MARs may represent the most accessible regions for integration in the DNA due to the open conformation and

high propensity for base-unpairing associated with the AT-richness. A third possibility relates to the observation that retroviruses may contain their own MARs. In fact, the mouse mammary tumor virus (MMTV) has been shown to contain a MAR in the LTR that binds a well characterized MAR-binding protein, SATB1 [50]. As the proviral pre-integration complex enters the nucleus, MAR binding proteins may bind and direct integration due to their affinity for binding to cellular MARs. It is known that sequence insertion within or near a MAR results in greatly reduced binding to the nuclear matrix [45]. In contrast, it has been shown that when retroviral integration occurs near MARs, contact with the nuclear matrix is maintained, suggesting that the presence of a MAR in the viral genome



**Figure 4**

Analysis of the relationship between integration near a MAR and integration within or between genes. The percentage of integrations that occurred within 2-kb of a MAR is reported for those that occurred within a gene or between genes. Data are reported for SL3-3 insertions at single insertion site (SL3-3 IS), SL3-3 insertions at common insertion sites (SL3-3 CIS), MLV, MoFe2, HTLV-1 and HIV-1.

may stabilize the contact between the chromosomal MAR and the nuclear matrix [49].

The selective advantage of integration near MARs may be that it positions the provirus in close proximity to transcription, RNA processing and transport machinery that is localized at the nuclear matrix (reviewed in [25]), thus activating expression from the viral promoter. In addition, our findings suggest the possibility that integration near MARs may have a role in malignant induction, specifically by gammaretroviruses. SL3-3 proviruses integrated at CISs in tumor DNA were shown to position preferentially within 2-kb upstream from a MAR, whereas SL3-3 proviruses integrated at single insertion sites in the same tumors did not show the same preference (Figure 3). Considering that gammaretroviruses like SL3-3 induce malignancy through insertional activation of oncogenes at CISs, this observation suggests that SL3-3 integration immediately upstream of MARs may be associated with activation of adjacent cellular gene expression. Such an effect might occur by disruption of the normal function of the MAR, thus altering local chromatin conformation. Changes in chromatin conformation, leading to changes in gene expression, are known to contribute to malignancy

(reviewed in [51]). Alternatively, integration at a specific distance and orientation with respect to a MAR may result in stimulation of expression from the viral promoter, thus enhancing virus-mediated activation of an adjacent cellular oncogene. Integration near MARs has also been implicated in malignant induction by small DNA tumor viruses [34]. These viruses do not induce disease by insertional activation; thus, the advantage of integration near MARs may relate to increased expression from the viral promoter.

Previous studies have reported that HIV-1, MLV, ASV and HTLV-1 prefer to integrate into genes [11,12,19,20]. In the present study, integration patterns of SL3-3 and MoFe2 were examined to determine if they also preferentially integrate into RefSeqs. Consistent with previous reports, our results indicated that SL3-3 proviruses at single insertion sites (40.3%) and MoFe2 proviruses (33.3%) integrate preferentially within RefSeqs as compared to the predicted frequency for random integrations (22%). SL3-3 proviruses integrated at CISs did not demonstrate the same preference, an observation consistent with the role of these integrants in enhancer-mediated activation of an adjacent oncogene. Of SL3-3 integrations at CISs that

occurred between genes, 68.8% were observed within 2-kb of a predicted MAR (Figure 4). Taken together, these studies provide additional evidence that proviral integration is not random, that MARs influence retroviral integration site selection, and that integration near MARs may have a role in the insertional activation of oncogenes by gammaretroviruses. Understanding the pressures that influence retroviral integration site selection is critical for further knowledge of the mechanisms of retroviral pathogenesis and for the development of retroviral vectors for gene-therapy.

## Methods

### Isolation of MoFe2-MuLV host-virus junction fragments

MoFe2 proviral integrations were analyzed from lymphomas induced in a previous study [40] and from acutely infected tissue culture cells. For that purpose,  $5 \times 10^5$  SC-1 murine fibroblasts at 25% confluence were infected with  $10^5$  infectious units (TCID<sub>50</sub>) of MoFe2 in the presence of 8 µg/ml of polybrene for 5 hours. Medium was removed, replaced with fresh EMEM with 10% FBS, and cells were harvested three days later. Genomic DNA was digested with DraI (TTT/AAA) or StuI (AGG/CCT), and libraries were constructed using Universal Genome Walker Kit (BD Biosciences) as described by the manufacturer. Libraries were constructed from both restriction enzyme digests to avoid introducing a bias for AT- or GC-rich sequences. Host-virus junction sequences were amplified by PCR using oligonucleotide primers and Universal Genome Walker Kit reagents as previously described [40]. Amplification products were cloned into TOPO-TA vector (Invitrogen Corp.) and submitted for automated sequence analysis. The resulting sequences were considered to represent valid MoFe2 integrations if they contained the viral 3' LTR and if the immediately flanking host sequence had a  $\geq 95\%$  identity to a single genomic locus.

### MAR analysis

A MAR prediction program termed MAR-Finder <http://www.futuresoft.org> was used to predict MARs on 60-kb intervals surrounding the insertion site using default detection and clipping parameters for SL3-3 (n = 86), MoFe2 (n = 42), MLV (n = 49), HIV-1 (n = 41) and HTLV-1 (n = 26) [34,41,42]. High scoring regions were considered valid if the average strength of a single peak representing a predicted MAR was  $>0.65$  [34].

### Competing interests

The author(s) declare that they have no competing interests.

### Authors' contributions

CNJ performed all experimental and computer-based analyses. LSL directed the experimental design, imple-

mentation and interpretation of data. Both authors read and approved the final manuscript.

## Acknowledgements

This work was supported by PHS grant CA83823, by Development Funds of the Tulane Cancer Center and by a grant from the Ladies Leukemia League. CNJ was supported in part by a grant from the Cancer Association of Greater New Orleans.

## References

1. Neil JC, Cameron ER: **Retroviral insertion sites and cancer: fountain of all knowledge?** *Cancer Cell* 2002, **2**(4):253-255.
2. Weidhaas JB, Angelichio EL, Fenner S, Coffin JM: **Relationship between retroviral DNA integration and gene expression.** *J Virol* 2000, **74**(18):8382-8389.
3. Robinson HL, Gagnon GC: **Patterns of proviral insertion and deletion in avian leukosis virus-induced lymphomas.** *J Virol* 1986, **57**(1):28-36.
4. Rohdewohld H, Weiher H, Reik W, Jaenisch R, Breindl M: **Retrovirus integration and chromatin structure: Moloney murine leukemia proviral integration sites map near DNase I-hypersensitive sites.** *J Virol* 1987, **61**(2):336-343.
5. Vijaya S, Steffen DL, Robinson HL: **Acceptor sites for retroviral integrations map near DNase I-hypersensitive sites in chromatin.** *J Virol* 1986, **60**(2):683-692.
6. Goodenow MM, Hayward WS: **5' long terminal repeats of myc-associated proviruses appear structurally intact but are functionally impaired in tumors induced by avian leukosis viruses.** *J Virol* 1987, **61**(8):2489-2498.
7. Panet A, Cedar H: **Selective degradation of integrated murine leukemia proviral DNA by deoxyribonucleases.** *Cell* 1977, **11**(4):933-940.
8. Leclercq I, Mortreux F, Gabet AS, Jonsson CB, Wattel E: **Basis of HTLV type I target site selection.** *AIDS Res Hum Retroviruses* 2000, **16**(16):1653-1659.
9. Scherding U, Rhodes K, Breindl M: **Transcriptionally active genome regions are preferred targets for retrovirus integration.** *J Virol* 1990, **64**(2):907-912.
10. Mooslehner K, Karls U, Harbers K: **Retroviral integration sites in transgenic Mov mice frequently map in the vicinity of transcribed DNA regions.** *J Virol* 1990, **64**(6):3056-3058.
11. Narezkina A, Taganov KD, Litwin S, Stoyanova R, Hayashi J, Seeger C, Skalka AM, Katz RA: **Genome-wide analyses of avian sarcoma virus integration sites.** *J Virol* 2004, **78**(21):11656-11663.
12. Ozawa T, Itoyama T, Sadamori N, Yamada Y, Hata T, Tomonaga M, Isobe M: **Rapid isolation of viral integration site reveals frequent integration of HTLV-I into expressed loci.** *J Hum Genet* 2004, **49**(3):154-165.
13. Stevens SW, Griffith JD: **Human immunodeficiency virus type I may preferentially integrate into chromatin occupied by LIHs repetitive elements.** *Proc Natl Acad Sci U S A* 1994, **91**(12):5557-5561.
14. Pruss D, Reeves R, Bushman FD, Wolffe AP: **The influence of DNA and nucleosome structure on integration events directed by HIV integrase.** *J Biol Chem* 1994, **269**(40):25031-25041.
15. Pruss D, Bushman FD, Wolffe AP: **Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core.** *Proc Natl Acad Sci U S A* 1994, **91**(13):5913-5917.
16. Pryciak PM, Varmus HE: **Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection.** *Cell* 1992, **69**(5):769-780.
17. Pryciak PM, Sil A, Varmus HE: **Retroviral integration into minichromosomes in vitro.** *Embo J* 1992, **11**(1):291-303.
18. Muller HP, Varmus HE: **DNA bending creates favored sites for retroviral integration: an explanation for preferred insertion sites in nucleosomes.** *Embo J* 1994, **13**(19):4704-4714.
19. Wu X, Li Y, Crise B, Burgess SM: **Transcription start regions in the human genome are favored targets for MLV integration.** *Science* 2003, **300**(5626):1749-1751.
20. Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F: **HIV-I integration in the human genome favors active genes and local hotspots.** *Cell* 2002, **110**(4):521-529.



21. Cereseto A, Giacca M: **Integration site selection by retroviruses.** *AIDS Rev* 2004, **6**(1):13-21.
22. Wu X, Li Y, Crise B, Burgess SM, Munroe DJ: **Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses.** *J Virol* 2005, **79**(8):5211-5214.
23. Wu X, Burgess SM: **Integration target site selection for retroviruses and transposable elements.** *Cell Mol Life Sci* 2004, **61**(19-20):2588-2596.
24. Holman AG, Coffin JM: **Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites.** *Proc Natl Acad Sci U S A* 2005, **102**(17):6103-6107.
25. Holth LT, Chadee DN, Spencer VA, Samuel SK, Safneck JR, Davie JR: **Chromatin, nuclear matrix and the cytoskeleton: role of cell structure in neoplastic transformation (review).** *Int J Oncol* 1998, **13**(4):827-837.
26. Bode J, Kohwi Y, Dickinson L, Joh T, Klehr D, Mielke C, Kohwi-Shigematsu T: **Biological significance of unwinding capability of nuclear matrix-associating DNAs.** *Science* 1992, **255**(5041):195-197.
27. Cockerill PN, Garrard WT: **Chromosomal loop anchorage sites appear to be evolutionarily conserved.** *FEBS Lett* 1986, **204**(1):5-7.
28. Yanagisawa J, Ando J, Nakayama J, Kohwi Y, Kohwi-Shigematsu T: **A matrix attachment region (MAR)-binding activity due to a p114 kilodalton protein is found only in human breast carcinomas and not in normal and benign breast disease tissues.** *Cancer Res* 1996, **56**(3):457-462.
29. Leman ES, Getzenberg RH: **Nuclear matrix proteins as biomarkers in prostate cancer.** *J Cell Biochem* 2002, **86**(2):213-223.
30. Coffey DS: **Nuclear matrix proteins as proteomic markers of preneoplastic and cancer lesions : commentary re: G. Brunagel et al., nuclear matrix protein alterations associated with colon cancer metastasis to the liver.** *Clin. Cancer Res.*, **8**: 3039-3045, 2002. *Clin Cancer Res* 2002, **8**(10):3031-3033.
31. Kurre P, Morris J, Thomasson B, Kohn DB, Kiem HP: **Scaffold attachment region-containing retrovirus vectors improve long-term proviral expression after transplantation of GFP-modified CD34+ baboon repopulating cells.** *Blood* 2003, **102**(9):3117-3119.
32. Agarwal M, Austin TW, Morel F, Chen J, Bohnlein E, Plavec I: **Scaffold attachment region-mediated enhancement of retroviral vector expression in primary T cells.** *J Virol* 1998, **72**(5):3720-3728.
33. Lutzko C, Senadheera D, Skelton D, Petersen D, Kohn DB: **Lentivirus vectors incorporating the immunoglobulin heavy chain enhancer and matrix attachment regions provide position-independent expression in B lymphocytes.** *J Virol* 2003, **77**(13):7341-7351.
34. Shera KA, Shera CA, McDougall JK: **Small tumor virus genomes are integrated near nuclear matrix attachment regions in transformed cells.** *J Virol* 2001, **75**(24):12339-12346.
35. Kulkarni A, Pavithra L, Rampalli S, Mogare D, Babu K, Shiekh G, Ghosh S, Chattopadhyay S: **HIV-1 integration sites are flanked by potential MARs that alone can act as promoters.** *Biochem Biophys Res Commun* 2004, **322**(2):672-677.
36. Hansen GM, Skapura D, Justice MJ: **Genetic profile of insertion mutations in mouse leukemias and lymphomas.** *Genome Res* 2000, **10**(2):237-243.
37. Li J, Shen H, Himmel KL, Dupuy AJ, Largaespada DA, Nakamura T, Shaughnessy JD, Jenkins NA, Copeland NG: **Leukaemia disease genes: large-scale cloning and pathway predictions.** *Nat Genet* 1999, **23**(3):348-353.
38. Mikkers H, Allen J, Knipscheer P, Romeijn L, Hart A, Vink E, Berns A, Romeyn L: **High-throughput retroviral tagging to identify components of specific signaling pathways in cancer.** *Nat Genet* 2002, **32**(1):153-159.
39. Starkey CR, Lobelle-Rich PA, Granger S, Brightman BK, Fan H, Levy LS: **Tumorigenic potential of a recombinant retrovirus containing sequences from Moloney murine leukemia virus and feline leukemia virus.** *J Virol* 1998, **72**(2):1078-1084.
40. Johnson C, Lobelle-Rich PA, Puetter A, Levy LS: **Substitution of feline leukemia virus long terminal repeat sequences into murine leukemia virus alters the pattern of insertional activation and identifies new common insertion sites.** *J Virol* 2005, **79**(1):57-66.
41. Glazko GV, Rogozin IB, Glazkov MV: **Comparative study and prediction of DNA fragments associated with various elements of the nuclear matrix.** *Biochim Biophys Acta* 2001, **1517**(3):351-364.
42. Rogozin IB, Glazko GV, Glazkov MV: **Computer prediction of sites associated with various elements of the nuclear matrix.** *Brief Bioinform* 2000, **1**(1):33-44.
43. Benham C, Kohwi-Shigematsu T, Bode J: **Stress-induced duplex DNA destabilization in scaffold/matrix attachment regions.** *J Mol Biol* 1997, **274**(2):181-196.
44. Albrecht P, Bode J, Buiting K, Prashanth AK, Lohmann DR: **Recurrent deletion of a region containing exon 24 of the RBI gene caused by non-homologous recombination between a LINE-1HS and MER21B element.** *J Med Genet* 2004, **41**(12):e122.
45. Goetze S, Gluch A, Benham C, Bode J: **Computational and in vitro analysis of destabilized DNA regions in the interferon gene cluster: potential of predicting functional gene domains.** *Biochemistry* 2003, **42**(1):154-166.
46. Goetze S, Baer A, Winkelmann S, Nehlsen K, Seibler J, Maass K, Bode J: **Performance of genomic bordering elements at predefined genomic loci.** *Mol Cell Biol* 2005, **25**(6):2260-2272.
47. Kim R, Trubetskoy A, Suzuki T, Jenkins NA, Copeland NG, Lenz J: **Genome-based identification of cancer genes by proviral tagging in mouse retrovirus-induced T-cell lymphomas.** *J Virol* 2003, **77**(3):2056-2062.
48. Klehr D, Maass K, Bode J: **Scaffold-attached regions from the human interferon beta domain can be used to enhance the stable expression of genes under the control of various promoters.** *Biochemistry* 1991, **30**(5):1264-1270.
49. Goetze S, Huesemann Y, Baer A, Bode J: **Functional characterization of transgene integration patterns by halo fluorescence in situ hybridization: electroporation versus retroviral infection.** *Biochemistry* 2003, **42**(23):7035-7043.
50. Liu J, Bramblett D, Zhu Q, Lozano M, Kobayashi R, Ross SR, Dudley JP: **The matrix attachment region-binding protein SATB1 participates in negative regulation of tissue-specific gene expression.** *Mol Cell Biol* 1997, **17**(9):5275-5287.
51. Zink D, Fischer AH, Nickerson JA: **Nuclear structure in cancer cells.** *Nat Rev Cancer* 2004, **4**(9):677-687.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
http://www.biomedcentral.com/info/publishing\_adv.asp

